

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 14, 2005, 04:20:34 ; Search time 4339.41 Seconds
(without alignments)
5024.367 Million cell updates/sec

Title: US-10-674-124A-1
Perfect score: 466
Sequence: 1 agatggatgatagatcatg.....actgaaacgtcgatattcag 466

Scoring table:  Gapop 60.0, Gapext 60.0

Searched: 41078325 seqs, 23393541228 residues

Word size: 0

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 50 summaries

Database : EST:
1: gb_est1:
2: gb_est2:
3: gb_est3:
4: gb_hc:
5: gb_est4:
6: gb_est5:
7: gb_est6:
8: gb_est7:
9: gb_gss1:
10: gb_gss2:
11: gb_gss3:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	268	57.5	406	5	EX488224 DKFZp686P
2	36	7.7	217	11	FR0012565
3	36	7.7	596	11	FR0012561
4	29	6.2	915	11	CNS04H62
5	29	6.2	1055	11	CNS03K08
6	28	6.0	261	2	BF910800
7	28	6.0	301	7	CV342994 MR0-UT004
8	28	6.0	618	10	CE424277
9	28	6.0	908	11	CR076016
10	27	5.8	491	11	CNS044H0
11	27	5.8	595	9	AQ602695
12	27	5.8	601	11	FR0012542
13	27	5.8	796	9	BZ149254
14	27	5.8	1101	11	CNS05PMW
15	26	5.6	547	11	CR254696
16	26	5.6	549	10	CE494613
17	26	5.6	578	9	BH258010
18	26	5.6	598	10	EX194414
19	26	5.6	678	10	CE586092
20	26	5.6	844	11	CNS042QD
21	26	5.6	968	11	CNS052LN
22	26	5.6	1027	10	CNS02PUN

23	25	5.4	237	9	BH055972
24	25	5.4	416	6	CB767615
25	25	5.4	492	9	AZ442057
26	25	5.4	498	6	CF608447
27	25	5.4	498	6	CF609035
28	25	5.4	498	6	CF609040
29	25	5.4	498	10	CE309451
30	25	5.4	547	9	AZ823003
31	25	5.4	576	9	AZ840510
32	25	5.4	641	9	BH091021
33	25	5.4	753	10	AG459390
34	25	5.4	902	11	CR105359
35	25	5.4	1030	11	CNS04T8S
36	25	5.4	1057	11	CNS05PZY
37	25	5.4	2998	4	AK043931
38	24	5.2	209	3	BM005463
39	24	5.2	231	3	BM026778
40	24	5.2	338	11	CR918692
41	24	5.2	398	10	CE767847
42	24	5.2	423	1	AV726011
43	24	5.2	433	9	AQ090767
44	24	5.2	434	9	BZ263771
45	24	5.2	498	10	CE725139
46	24	5.2	526	10	CE822229
47	24	5.2	567	9	AZ271663
48	24	5.2	607	9	AZ624964
49	24	5.2	620	10	AG000879
50	24	5.2	622	10	AG000876

ALIGNMENTS

RESULT 1

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

406 bp mRNA linear EST 04-SEP-2003
DKFZp686P23268_r1 686 (synonym: hlcc3) Homo sapiens cDNA clone
DKFZp686P23268_5', mRNA sequence.
BX488224
BX488224
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 406)
Bahr, A., Lauber, J., Mewes, H.W., Weil, B., Amid, C., Osanger, A.,
Fobo, G., Han, M. and Wiemann, S.
EST (Bahr, A., Lauber, J., Mewes, H.W., Weil, B., et al.)
Unpublished (2003)
Contact: MIPS
MIPS
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by Qiagen (Hilden/Germany) within the cDNA sequencing
consortium of the German Genome Project.
No si sequence available.
This clone (DKFZp686P23268) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES

source

Location/Qualifiers

1..406

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="DKFZp686P23268"

/dev_stage="adult"

/lab_host="DH10B"

/clone_lib="686 (synonym: hlcc3)"

/note="Vector: pTriplex2; Site_1: SfiIA; Site_2: SfiIB;

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 14, 2005, 03:13:04 ; Search time 515.162 Seconds
(without alignments)
6028.680 Million cell updates/sec

Title: US-10-674-124A-1
Perfect score: 466
Sequence: 1 agatgatgatagatcattg.....actgaaacgtcggaatttcag 466

Scoring table: OLIGO NUC
Gapop 60.0, Gapext 60.0

Searched: 4996997 seqs, 332346308 residues

Word size: 0

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 50 summaries

Database: N_Geneseq_21.*
1: Geneseq1980s.*
2: Geneseq1990s.*
3: Geneseq2000s.*
4: Geneseq2001as.*
5: Geneseq2001bs.*
6: Geneseq2002as.*
7: Geneseq2002bs.*
8: Geneseq2003as.*
9: Geneseq2003bs.*
10: Geneseq2003cs.*
11: Geneseq2003ds.*
12: Geneseq2004as.*
13: Geneseq2004bs.*
14: Geneseq2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	31	6.7	76644	12 ADQ97602	Adq97602 Mouse can
2	26	5.6	151858	13 ABD33489	Abd33489 Murine ca
3	25	5.4	8222	8 ACF62815	Acf62815 Colon can
4	25	5.4	8222	8 ACF62793	Acf62793 Colon can
5	25	5.4	11222	10 ADB54189	Adb54189 Pretreat
6	25	5.4	11222	10 ADB54317	Adb54317 Pretreat
7	24	5.2	393	4 AAI87991	Aai87991 Human pol
8	24	5.2	4372	10 ADD48700	Add48700 Rat gene
9	24	5.2	24281	13 ABD33480	Abd33480 Murine ca
10	24	5.2	24281	14 AD213479	Ad213479 Murine ca
11	24	5.2	95835	11 ACN45042	Acn45042 Human gen
12	24	5.2	110000	13 ABD32804_4	Continuation (5 of
13	23	4.9	551	6 ABN60833	Abn60833 Human can
14	23	4.9	6731	6 ABK39962	Abk39962 Human che
15	23	4.9	85920	14 AD213418	Ad213418 Human can
16	22	4.7	355	2 AAX33665	Aax33665 DNA cand
17	22	4.7	397	6 ABN79099	Abn79099 Human ORF
18	22	4.7	519	12 ACH76810	Ach76810 Human gen
19	22	4.7	26502	12 ADL18577	Adl18577 Human wil

C	20	22	4.7	39071	12	ADM97420	Adm97420 Prostate
C	21	22	4.7	42952	13	ABD33609	Abd33609 Murine ca
C	22	22	4.7	46275	6	ABT10145	Abt10145 Human bre
C	23	22	4.7	46275	10	ADL13621	Adl13621 Osteoarth
C	24	22	4.7	88232	11	ACN44980	Acn44980 Mouse gen
C	25	22	4.7	96596	10	ADE95968	Ade95968 Human NFA
C	26	22	4.7	96597	9	ADA02720	Ada02720 Human NFA
C	27	22	4.7	96597	10	ADB72458	Adb72458 Human NFA
C	28	22	4.7	96960	8	ACF62734	Acf62734 Cancer ba
C	29	22	4.7	96960	8	ADB20849	Adb20849 MRPI base
C	30	22	4.7	96960	10	ADB87938	Adb87938 Human UGT
C	31	22	4.7	96960	10	ADB96921	Adb96921 Human UGT
C	32	22	4.7	96960	10	ADB92112	Adb92112 Human MDR
C	33	22	4.7	100534	13	ABD32647	Abd32647 Mouse can
C	34	22	4.7	107432	12	ADQ97161	Adq97161 Mouse can
C	35	22	4.7	114771	12	ADQ17641	Adq17641 Human sof
C	36	22	4.7	123785	10	ABX77171	Abx77171 DNA seque
C	37	22	4.7	158001	12	ADL17884	Adl17884 Human pho
C	38	22	4.7	165156	13	ADS36459	Ads36459 Human aut
C	39	22	4.7	173564	13	ABD32953	Abd32953 Human can
C	40	22	4.7	183610	8	ACF62736	Acf62736 Cancer ba
C	41	22	4.7	183610	8	ADB20851	Adb20851 MRPI base
C	42	22	4.7	183610	10	ADB87940	Adb87940 Human UGT
C	43	22	4.7	183610	10	ADB96923	Adb96923 Human MDR
C	44	22	4.7	183610	10	ADB92114	Adb92114 Human MDR
C	45	22	4.7	186591	8	ACF62750	Acf62750 Cancer ba
C	46	22	4.7	186591	8	ADB20869	Adb20869 MRPI base
C	47	22	4.7	186591	10	ADB87958	Adb87958 Human UGT
C	48	22	4.7	186591	10	ADB96941	Adb96941 Human MDR
C	49	22	4.7	186591	10	ADB92132	Adb92132 Human MDR
C	50	22	4.7	208648	8	ACF62735	Acf62735 Cancer ba

ALIGNMENTS

RESULT 1
ADQ97602
ID ADQ97602 standard; DNA; 76644 BP.
XX
AC ADQ97602;
XX
DT 07-OCT-2004 (first entry)
XX
DE Mouse cancer associated sequence MD10-013, SEQ ID 579.
XX
KW Cytostatic; Gene Therapy; cancer; leukemia; lymphoma; Mouse; ds.
XX
OS Mus musculus.
XX
PN WO2004060304-A2.
XX
PD 22-JUL-2004.
XX
PF 22-DEC-2003; 2003WO-US041389.
XX
PR 27-DEC-2002; 2002US-00330773.
XX
PA (SAGR-) SAGRES DISCOVERY INC.
PI Morris DW, Malandro MS;
XX
WPI; 2004-543781/52.
XX
PT New isolated cancer associated nucleic acids comprising at least 10 contiguous nucleotides, useful for diagnosing, preventing and/or treating cancers such as leukemia and lymphoma.
XX
PS Claim 1; SEQ ID NO 579; 199pp; English.
XX
CC The present invention relates to cancer associated sequences (ADQ97025-ADQ98004). The sequences are useful for the diagnosis, prevention and/or treatment of cancer, such as leukemia and lymphoma. Note: The sequence


```
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 175240
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-175240

Query Match          5.4%; Score 25; DB 3; Length 601;
Best Local Similarity 100.0%; Pred. No. 0.014;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 93 TGATTGATAGTAGGTAGTAGTAG 117
Db 559 TGATTGATAGTAGGTAGTAGTAG 583

RESULT 3
US-09-949-016-175241
; Sequence 175241, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 175241
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-175241

Query Match          5.4%; Score 25; DB 3; Length 601;
Best Local Similarity 100.0%; Pred. No. 0.014;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 93 TGATTGATAGTAGGTAGTAGTAG 117
Db 365 TGATTGATAGTAGGTAGTAGTAG 389

RESULT 4
US-09-949-016-16717
; Sequence 16717, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
```

```
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16717
; LENGTH: 20662
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-16717

Query Match          5.4%; Score 25; DB 3; Length 20662;
Best Local Similarity 100.0%; Pred. No. 0.013;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 93 TGATTGATAGTAGGTAGTAGTAG 117
Db 6351 TGATTGATAGTAGGTAGTAGTAG 6375

RESULT 5
5175383-1/c
; Patent No. 5175383
; APPLICANT: LEDER, PHILIP;MULLER, WILLIAM J.
; TITLE OF INVENTION: ANIMAL MODEL FOR BENIGN PROSTATIC DISEASE
; NUMBER OF SEQUENCES: 9
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/312,641
; FILING DATE: 17-FEB-1989
; SEQ ID NO:1
; LENGTH: 7868
5175383-1

Query Match          5.2%; Score 24; DB 9; Length 7868;
Best Local Similarity 100.0%; Pred. No. 0.039;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 206 GATAGATGGATGGATGGATGGATAG 229
Db 5087 GATAGATGGATGGATGGATGGATAG 5064

RESULT 6
US-09-949-016-15616
; Sequence 15616, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15616
; LENGTH: 59719
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(59719)
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 14, 2005, 06:01:45 : Search time 361.444 Seconds
(without alignments)
631.944 Million cell updates/sec

Title: US-10-674-124A-1
Perfect score: 466
Sequence: 1 agatgatgatgatcattg.....actgaacgtcgatttcag 466

Scoring table: OLIGO NUC
Gapop 60.0 , Gapext 60.0

Searched: 4161359 seqs, 245077644 residues

Word size : 0
Total number of hits satisfying chosen parameters: 8322718

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 50 summaries

Database : Published Applications NA New:
1: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
2: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	22	4.7	165156	6	US-10-995-561-13304
2	21	4.5	201	6	US-10-995-561-20634
3	21	4.5	1548	6	US-10-750-185-34790
4	21	4.5	13943	9	US-11-012-762-39
5	21	4.5	44718	6	US-10-995-561-13217
6	21	4.5	169495	7	US-11-121-086-61
7	21	4.5	169725	7	US-11-121-086-63
8	21	4.5	200628	7	US-11-121-086-62
9	20	4.3	201	6	US-10-995-561-20635
10	20	4.3	201	6	US-10-995-561-23773
11	20	4.3	201	6	US-10-995-561-28163
12	20	4.3	201	6	US-10-995-561-30516
13	20	4.3	201	6	US-10-995-561-47793
14	20	4.3	201	6	US-10-995-561-47795
15	20	4.3	201	6	US-10-995-561-47798
16	20	4.3	201	6	US-10-995-561-47801
17	20	4.3	201	6	US-10-995-561-47804
18	20	4.3	201	6	US-10-995-561-47808
19	20	4.3	201	6	US-10-995-561-47812
20	20	4.3	201	6	US-10-995-561-47816
21	20	4.3	201	6	US-10-995-561-77732
22	20	4.3	1357	6	US-10-750-185-57399
23	20	4.3	1413	6	US-10-750-185-57194

C 24	4.3	1632	6	US-10-750-185-58027	Sequence 58027, A
C 25	4.3	1935	6	US-10-750-185-51171	Sequence 51171, A
C 26	4.3	2103	6	US-10-750-185-29643	Sequence 29643, A
C 27	4.3	2129	7	US-11-08-528-33	Sequence 33, Appl
C 28	4.3	2231	6	US-10-750-185-34561	Sequence 34561, A
C 29	4.3	31973	6	US-10-995-561-13229	Sequence 13229, A
C 30	4.3	52192	6	US-10-995-561-13231	Sequence 13231, A
C 31	4.3	110847	7	US-11-121-086-11	Sequence 11, Appl
C 32	4.3	119160	7	US-11-121-086-56	Sequence 56, Appl
C 33	4.3	150468	7	US-11-121-086-59	Sequence 59, Appl
C 34	4.3	152637	7	US-11-121-086-55	Sequence 55, Appl
C 35	4.3	175673	7	US-11-121-086-79	Sequence 79, Appl
C 36	4.3	177175	7	US-11-121-086-79	Sequence 79, Appl
C 37	4.3	180654	7	US-11-121-086-58	Sequence 58, Appl
C 38	4.3	18786	6	US-10-995-561-13474	Sequence 13474, A
C 39	4.3	187986	6	US-10-995-561-13252	Sequence 13252, A
C 40	4.3	193789	7	US-11-112-908-55	Sequence 55, Appl
C 41	4.3	222094	6	US-10-995-561-13244	Sequence 13244, A
C 42	4.3	285300	6	US-10-857-780-6	Sequence 6, Appl
C 43	4.3	645179	6	US-10-995-561-13293	Sequence 13293, A
C 44	4.3	1125000	6	US-10-995-561-13286	Sequence 13286, A
C 45	4.1	201	6	US-10-995-561-23615	Sequence 23615, A
C 46	4.1	201	6	US-10-995-561-23805	Sequence 23805, A
C 47	4.1	201	6	US-10-995-561-43973	Sequence 43973, A
C 48	4.1	201	6	US-10-995-561-43974	Sequence 43974, A
C 49	4.1	201	6	US-10-995-561-44636	Sequence 44636, A
C 50	4.1	201	6	US-10-995-561-44637	Sequence 44637, A

ALIGNMENTS

RESULT 1
US-10-995-561-13304
; Sequence 13304, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13304
; LENGTH: 165156
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-13304
Query Match 4.7%; Score 22; DB 6; Length 165156;
Best Local Similarity 100.0%; Pred. No. 0.062;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 206 GATAGTGGATGGATGGATGAT 227
DB 124971 GATAGTGGATGGATGGATGAT 124992
RESULT 2
US-10-995-561-20634
; Sequence 20634, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24

```
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20634
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-20634

Query Match      4.5%; Score 21; DB 6; Length 201;
Best Local Similarity 100.0%; Pred. No. 0.16;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 328 GTAGATAGATGATGATGGAT 348
Db 76 GTAGATAGATGATGATGGAT 96

RESULT 3
US-10-750-185-34790/c
; Sequence 34790, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM11100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 34790
; LENGTH: 1548
; TYPE: DNA
; ORGANISM: Bovine 19866880504043
US-10-750-185-34790

Query Match      4.5%; Score 21; DB 6; Length 1548;
Best Local Similarity 100.0%; Pred. No. 0.17;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 206 GATAGATGGATGGATGGATGA 226
Db 1198 GATAGATGGATGGATGGATGA 1178

RESULT 4
US-11-012-762-39/c
; Sequence 39, Application US/11012762
; Publication No. US20050244815A1
; GENERAL INFORMATION:
; APPLICANT: Georgia State University Research Foundation, Inc.
; TITLE OF INVENTION: Compositions and Methods for Viral Resistance Genes
; FILE REFERENCE: GSU1.PCT
; CURRENT APPLICATION NUMBER: US/11/012,762
; CURRENT FILING DATE: 2004-12-15
; PRIOR APPLICATION NUMBER: PCT/US03/19300
; PRIOR FILING DATE: 2003-06-19
; PRIOR APPLICATION NUMBER: US 60/390,046
; PRIOR FILING DATE: 2002-06-19
; NUMBER OF SEQ ID NOS: 133
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 39
; LENGTH: 13943
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
```

```
; NAME/KEY: CDS
; LOCATION: (1514)..(1693)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3240)..(3516)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (6428)..(6612)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (10018)..(10250)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (10901)..(11054)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (12252)..(12350)
; US-11-012-762-39

Query Match      4.5%; Score 21; DB 9; Length 13943;
Best Local Similarity 100.0%; Pred. No. 0.18;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 205 TGATAGATGGATGGATGGATG 225
Db 3155 TGATAGATGGATGGATGGATG 3135

RESULT 5
US-10-995-561-13217
; Sequence 13217, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13217
; LENGTH: 44718
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-13217

Query Match      4.5%; Score 21; DB 6; Length 44718;
Best Local Similarity 100.0%; Pred. No. 0.19;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 328 GTAGATAGATGATGATGGAT 348
Db 26622 GTAGATAGATGATGATGGAT 26642

RESULT 6
US-11-121-086-61
; Sequence 61, Application US/11121086
; Publication No. US20050286459A1
; GENERAL INFORMATION:
; APPLICANT: POULSEN, TIM S.
; APPLICANT: NIELSEN, KIRSTEN V.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138.6000-00000
; CURRENT APPLICATION NUMBER: US/11/121,086
; CURRENT FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,570
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 61
```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 13, 2005, 22:13:37 ; Search time 2708.06 Seconds
(without alignments)
9781.550 Million cell updates/sec

Title: US-10-674-124A-1
Perfect score: 466
Sequence: 1 agatggatgatgatcattg.....actgaaacgtcgaatttcag 466

Scoring table: OLIGO-NUC
Gapop 60.0, Gapext 60.0

Searched: 5883141 seqs, 28421725653 residues

Word size : 0
Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 50 summaries

Database : GenEmbl.*
1: gb.ba.*
2: gb.in.*
3: gb.env.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pr.*
9: gb.ro.*
10: gb.sts.*
11: gb.sv.*
12: gb.un.*
13: gb.vi.*
14: gb.hcg.*
15: gb.pl.*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	466	100.0	144432	14 AF276983	AF276983 Homo sapi
C 2	466	100.0	179567	8 AL669831	AL669831 Human DNA
C 3	244	52.4	701	10 HUMUT7925	L30383 Human STS U
C 4	92	19.7	389	10 AB144915	AB144915 Homo sapi
C 5	92	19.7	85404	8 AC008165	AC008165 Homo sapi
C 6	92	19.7	169754	8 AC103954	AC103954 Homo sapi
C 7	92	19.7	175263	8 AC136777	AC136777 Homo sapi
C 8	65	13.9	70720	14 AC027823	AC027823 Homo sapi
C 9	57	12.2	148535	14 AF166490	AF166490 Homo sapi
C 10	32	6.9	173965	9 AC135673	AC135673 Mus muscu
C 11	32	6.9	204505	14 AC119344	AC119344 Rattus no
C 12	31	6.7	68336	14 AC100653	AC100653 Mus muscu
C 13	31	6.7	110000	14 AC151851_0	AC151851 Takifugu
C 14	31	6.7	112785	9 AL929332	AL929332 Mouse DNA
C 15	31	6.7	125681	9 AC122889	AC122889 Mus muscu
C 16	31	6.7	142195	5 AL844597	AL844597 Zebrafish
C 17	31	6.7	159743	9 AC153971	AC153971 Mus muscu
C 18	31	6.7	164582	9 AC153899	AC153899 Mus muscu

C 19	31	6.7	185068	9 AC157093	AC157093 Mus muscu
C 20	31	6.7	213959	14 AC102642	AC102642 Mus muscu
C 21	31	6.7	228313	14 AC097557	AC097557 Rattus no
C 22	31	6.7	231928	14 AC135680	AC135680 Rattus no
C 23	31	6.7	232002	9 AC153587	AC153587 Mus muscu
C 24	31	6.7	233830	9 AC122223	AC122223 Mus muscu
C 25	31	6.7	241738	14 AC094839	AC094839 Rattus no
C 26	31	6.7	244928	14 AC097177	AC097177 Rattus no
C 27	31	6.7	245125	14 AC133759	AC133759 Rattus no
C 28	31	6.7	246646	14 AC095667	AC095667 Rattus no
C 29	31	6.7	247269	14 AC120819	AC120819 Rattus no
C 30	31	6.7	251286	14 AC125835	AC125835 Rattus no
C 31	31	6.7	272329	14 AC094789	AC094789 Rattus no
C 32	30	6.4	86719	8 HSJ906P16	AL079339 Human DNA
C 33	30	6.4	163243	8 AL513304	AL513304 Human DNA
C 34	30	6.4	213624	14 AC129630	AC129630 Rattus no
C 35	30	6.4	238725	14 AC115260	AC115260 Rattus no
C 36	30	6.4	239049	14 AC111870	AC111870 Rattus no
C 37	29	6.2	74149	14 AC123879_3	Continuation (4 of
C 38	29	6.2	76522	14 AC160993	AC160993 Mus muscu
C 39	29	6.2	110000	14 AC123648_0	AC123648 Mus muscu
C 40	29	6.2	125651	14 CR376840	CR376840 Danio rer
C 41	29	6.2	131206	14 AC154449	AC154449 Mus muscu
C 42	29	6.2	150925	8 AC113607	AC113607 Homo sapi
C 43	29	6.2	150925	8 AC125282	AC125282 Homo sapi
C 44	29	6.2	189169	14 AC122690	AC122690 Homo sapi
C 45	29	6.2	211422	8 AC073493	AC073493 Homo sapi
C 46	28	6.0	110000	14 AC008576_0	AC008576 Homo sapi
C 47	28	6.0	110000	14 AC151851_1	Continuation (2 of
C 48	28	6.0	111173	8 AL513204	AL513204 Human DNA
C 49	28	6.0	116180	14 AC133296	AC133296 Rattus no
C 50	28	6.0	117077	8 HSJ662M14	AL079336 Human DNA

ALIGNMENTS

AF276983 144432 bp DNA linear HTG 26-JUL-2002
Homo sapiens chromosome 8 clone RP11-112904, WORKING DRAFT
SEQUENCE, 45 unordered pieces.

AF276983
HTG; HTGS PHASE1; HTGS_DRAFT; HTGS_CANCELLED.
Homo sapiens (human)

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.

1 (bases 1 to 144432)
Schilhabel,M.B., Baumgart,C., Blechschmidt,K., Dette,M., Jahn,N.,
Lehmann,R., Menzel,U., Polley,A., Reichwald,K., Schudy,A.,
Siddiqui,R., Taudien,S., Wen,G., Rosenthal,A. and Platzzer,M.

Chromosome 8 genomic sequence

Unpublished

2 (bases 1 to 144432)

Genome Sequencing Center Jena.

Direct Submission

Submitted (09-JUN-2000) Genome Analysis, Institute of Molecular

Biotechnology, Beutenbergstr. 11, Jena 07745, Germany

On Jun 7, 2001 this sequence version replaced gi:8575834.

----- Genome Center

Center: Institute of Molecular Biotechnology

Center code: IMB

Web site: <http://genome.imb-jena.de/>

Contact: gsdj-submit@genome.imb-jena.de

----- Project Information

Center project name: H502

Center clone name: RP11-112904

----- Summary Statistics

Sequencing vector: M13; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.990329
 Consensus quality: 112528 bases at least Q40
 Consensus quality: 121546 bases at least Q30
 Consensus quality: 129443 bases at least Q20
 Quality coverage: 3,46 in Q20 bases; sum-of-contigs

 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 45 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence.
 * as soon as it is available and the accession number will
 * be preserved.

```

1
* 4718: contig of 4718 bp in length
* 4819: gap of unknown length
* 4819: contig of 14579 bp in length
* 19398: gap of unknown length
* 19498: contig of 11312 bp in length
* 30810: gap of unknown length
* 30909: contig of 3029 bp in length
* 33939: gap of unknown length
* 34038: contig of 2028 bp in length
* 34039: gap of unknown length
* 36067: contig of 10196 bp in length
* 36167: gap of unknown length
* 46362: contig of 10196 bp in length
* 46363: gap of unknown length
* 46463: contig of 8629 bp in length
* 55091: gap of unknown length
* 55092: contig of 5265 bp in length
* 55191: gap of unknown length
* 60456: gap of unknown length
* 60457: contig of 6019 bp in length
* 60557: gap of unknown length
* 66675: gap of unknown length
* 66676: contig of 3097 bp in length
* 69772: gap of unknown length
* 69873: contig of 6680 bp in length
* 76552: gap of unknown length
* 76553: contig of 3535 bp in length
* 80187: gap of unknown length
* 80287: contig of 2732 bp in length
* 83019: gap of unknown length
* 83119: contig of 3516 bp in length
* 86635: gap of unknown length
* 86735: contig of 4557 bp in length
* 91292: gap of unknown length
* 91293: contig of 2039 bp in length
* 93431: gap of unknown length
* 93432: contig of 5515 bp in length
* 93532: gap of unknown length
* 99047: contig of 3395 bp in length
* 102541: gap of unknown length
* 102542: contig of 1441 bp in length
* 104082: gap of unknown length
* 104182: contig of 1214 bp in length
* 105396: gap of unknown length
* 105496: contig of 1469 bp in length
* 105965: gap of unknown length
* 106966: contig of 1045 bp in length
* 107066: gap of unknown length
* 108110: contig of 1482 bp in length
* 108111: gap of unknown length
* 109692: contig of 1504 bp in length
* 109792: gap of unknown length
* 111296: contig of 1279 bp in length
* 111297: gap of unknown length
* 111397: contig of 1364 bp in length
* 112760: gap of unknown length
* 112860: contig of 1404 bp in length
* 112861: gap of unknown length
* 114264: contig of 1279 bp in length
* 114265: gap of unknown length
* 115643: contig of 1215 bp in length
* 115644: gap of unknown length
* 117543: contig of 1420 bp in length
* 117544: gap of unknown length
* 117058: contig of 1333 bp in length
* 117059: gap of unknown length
* 118478: contig of 1333 bp in length
* 118479: gap of unknown length
* 118579: contig of 1333 bp in length

```

```

* 119912 120011: gap of unknown length
* 120012 121916: contig of 1905 bp in length
* 121917 122016: gap of unknown length
* 122017 12282: contig of 1266 bp in length
* 123283 123382: gap of unknown length
* 123383 124842: contig of 1460 bp in length
* 124843 124942: gap of unknown length
* 124943 126731: contig of 1789 bp in length
* 126732 126831: gap of unknown length
* 126832 128210: contig of 1379 bp in length
* 128211 128310: gap of unknown length
* 128311 129577: contig of 1267 bp in length
* 129578 129677: gap of unknown length
* 129679 131529: contig of 1852 bp in length
* 131530 131629: gap of unknown length
* 131630 132683: contig of 1054 bp in length
* 132684 132783: gap of unknown length
* 132784 134301: contig of 1518 bp in length
* 134302 134401: gap of unknown length
* 134402 136280: contig of 1859 bp in length
* 136281 136360: gap of unknown length
* 136361 137553: contig of 1193 bp in length
* 137554 137653: gap of unknown length
* 137654 139508: contig of 1855 bp in length
* 139509 139608: gap of unknown length
* 139609 141230: contig of 1622 bp in length
* 141231 141330: gap of unknown length
* 141331 143051: contig of 1721 bp in length
* 143052 143151: gap of unknown length
* 143152 144432: contig of 1281 bp in length.

FEATURES
      source
1. .144432
   /organism="Homo sapiens"
   /mol_type="genomic DNA"
   /db_xref="taxon:9606"
   /chromosome="8"
   /clone="RP11-112904"
   4719. .4818
   /estimated_length=unknown
   19398. .19497
   /estimated_length=unknown
   30810. .30909
   /estimated_length=unknown
   33939. .34038
   /estimated_length=unknown
   36067. .36166
   /estimated_length=unknown
   46363. .46462
   /estimated_length=unknown
   55092. .55191
   /estimated_length=unknown
   60457. .60556
   /estimated_length=unknown
   66576. .66675
   /estimated_length=unknown
   69773. .69872
   /estimated_length=unknown
   76553. .76652
   /estimated_length=unknown
   80188. .80287
   /estimated_length=unknown
   83020. .83119
   /estimated_length=unknown
   86636. .86735
   /estimated_length=unknown
   91293. .91392
   /estimated_length=unknown
   93432. .93531
   /estimated_length=unknown
   99047. .99146
   /estimated_length=unknown
   102542. .102641
   /estimated_length=unknown

```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 14, 2005, 06:01:45 ; Search time 160.556 Seconds
(without alignments)
631.944 Million cell updates/sec

Title: US-10-674-124A-2

Perfect score: 207

Sequence: 1 aaccagcttagcagacgc.....agcgggtaccagactcaat 207

Scoring table: POLYGNUG

Gapop 60.0 , Gapext 60.0

Searched: 4161359 seqs, 245077644 residues

Word size: 10

Total number of hits satisfying chosen parameters: 8322718

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 50 summaries

Database : Published Applications NA New:*

- 1: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
- 2: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
- 5: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
- 7: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
- 9: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
- 10: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	30	14.5	169047	7	US-11-121-086-15
2	29	14.0	2443	6	US-10-750-185-40370
3	27	13.0	201	6	US-10-995-561-33632
4	27	13.0	201	6	US-10-995-561-33668
5	27	13.0	201	6	US-10-995-561-76924
6	27	13.0	201	6	US-10-995-561-76927
7	27	13.0	201	6	US-10-995-561-77405
8	27	13.0	693	6	US-10-750-185-31742
9	27	13.0	1338	6	US-10-750-185-58661
10	27	13.0	1879	6	US-10-750-185-38372
11	27	13.0	1968	6	US-10-750-185-24617
12	27	13.0	2090	6	US-10-750-185-51962
13	27	13.0	3098	6	US-10-750-185-46414
14	27	13.0	3238	6	US-10-750-185-38644
15	27	13.0	110711	6	US-10-995-561-13264
16	27	13.0	147700	6	US-10-857-780-3
17	27	13.0	176503	7	US-11-121-086-53
18	27	13.0	394468	6	US-10-995-561-13473
19	26	12.6	600	6	US-10-750-185-3422
20	26	12.6	888	6	US-10-750-185-35928
21	26	12.6	1085	6	US-10-750-185-28972
22	26	12.6	1089	6	US-10-750-185-40317
23	26	12.6	1488	6	US-10-750-185-60015

C	24	26	12.6	1500	6	US-10-750-185-24893	Sequence 24893, A
C	25	26	12.6	1626	6	US-10-750-185-42949	Sequence 42949, A
C	26	26	12.6	1768	6	US-10-750-185-48978	Sequence 48978, A
C	27	26	12.6	1806	6	US-10-750-185-49853	Sequence 49853, A
C	28	26	12.6	1911	6	US-10-750-185-42482	Sequence 42482, A
C	29	26	12.6	1973	6	US-10-750-185-64418	Sequence 64418, A
C	30	26	12.6	2174	6	US-10-750-185-31781	Sequence 31781, A
C	31	26	12.6	2386	6	US-10-750-185-28178	Sequence 28178, A
C	32	26	12.6	2822	6	US-10-750-185-28046	Sequence 28046, A
C	33	26	12.6	2859	6	US-10-750-185-60847	Sequence 60847, A
C	34	26	12.6	3015	6	US-10-750-185-48995	Sequence 48995, A
C	35	26	12.6	3336	6	US-10-750-185-45825	Sequence 45825, A
C	36	26	12.6	3581	6	US-10-750-185-55045	Sequence 55045, A
C	37	26	12.6	56448	6	US-10-995-561-13359	Sequence 13359, A
C	38	26	12.6	60844	6	US-10-995-561-13359	Sequence 13359, A
C	39	26	12.6	66131	6	US-10-995-561-13501	Sequence 13501, A
C	40	26	12.6	119160	7	US-11-121-086-12	Sequence 12, Appl
C	41	26	12.6	179666	7	US-11-121-086-67	Sequence 67, Appl
C	42	26	12.6	207600	7	US-11-112-908-31	Sequence 31, Appl
C	43	26	12.6	403278	6	US-10-995-561-13421	Sequence 13421, A
C	44	25	12.1	201	6	US-10-995-561-9503	Sequence 9503, Ap
C	45	25	12.1	201	6	US-10-995-561-9519	Sequence 9519, Ap
C	46	25	12.1	201	6	US-10-995-561-51200	Sequence 51200, A
C	47	25	12.1	201	6	US-10-995-561-53066	Sequence 53066, A
C	48	25	12.1	201	6	US-10-995-561-55588	Sequence 55588, A
C	49	25	12.1	201	6	US-10-995-561-75984	Sequence 75984, A
C	50	25	12.1	766	6	US-10-750-185-55672	Sequence 55672, A

ALIGNMENTS

RESULT 1
US-11-121-086-15
; Sequence 15, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
; APPLICANT: POULSEN, TIM S.
; APPLICANT: NIELSEN, KIRSTEN V.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138.6000-00000
; CURRENT APPLICATION NUMBER: US/11/121,086
; CURRENT FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,570
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 15
; LENGTH: 169047
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-121-086-15

Query Match 14.5%; Score 30; DB 7; Length 169047;
Best Local Similarity 100.0%; Pred. No. 4.2e-07;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 CACACACACACATGTCACGCACACACA 80
Db 39337 CACACACACACATGTCACGCACACACA 39366

RESULT 2
US-10-750-185-40370
; Sequence 40370, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: Denise, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen

```
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM11100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 40370
; LENGTH: 2443
; TYPE: DNA
; ORGANISM: Bovine 19866880503606
US-10-750-185-40370

Query Match 14.0%; Score 29; DB 6; Length 2443;
Best Local Similarity 100.0%; Pred. No. 1.9e-06;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 36 AGGCACACGCACACACACACACACACA 64
Db 2396 AGGCACACGCACACACACACACACA 2424

RESULT 3
US-10-995-561-33632
; Sequence 33632, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 33632
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-33632

Query Match 13.0%; Score 27; DB 6; Length 201;
Best Local Similarity 100.0%; Pred. No. 2.7e-05;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 45 CACACACACACACACACATGCACAC 71
Db 113 CACACACACACACACACATGCACAC 139

RESULT 4
US-10-995-561-33668
; Sequence 33668, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 33668
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-33668

Query Match 13.0%; Score 27; DB 6; Length 201;
Best Local Similarity 100.0%; Pred. No. 2.7e-05;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 45 CACACACACACACACACATGCACAC 71
Db 104 CACACACACACACACACATGCACAC 130

RESULT 7
US-10-995-561-77405
; Sequence 77405, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
```

```
Query Match 13.0%; Score 27; DB 6; Length 201;
Best Local Similarity 100.0%; Pred. No. 2.7e-05;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 45 CACACACACACACACACATGCACAC 71
Db 109 CACACACACACACACACATGCACAC 135

RESULT 5
US-10-995-561-76924
; Sequence 76924, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 76924
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-76924

Query Match 13.0%; Score 27; DB 6; Length 201;
Best Local Similarity 100.0%; Pred. No. 2.7e-05;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 45 CACACACACACACACACATGCACAC 71
Db 169 CACACACACACACACACATGCACAC 195

RESULT 6
US-10-995-561-76927
; Sequence 76927, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 76927
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-76927

Query Match 13.0%; Score 27; DB 6; Length 201;
Best Local Similarity 100.0%; Pred. No. 2.7e-05;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 45 CACACACACACACACACATGCACAC 71
Db 104 CACACACACACACACACATGCACAC 130

RESULT 7
US-10-995-561-77405
; Sequence 77405, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 14, 2005, 04:54:42 : Search time 414.923 Seconds
(without alignments)
4125.494 Million cell updates/sec

Title: US-10-674-124A-2
Perfect score: 207
Sequence: 1 aacagcgttagcagacgc.....agcgggtaccagagctcaat 207

Scoring table: OLIGO NUC
Gapop 60.0, Gapext 60.0

Searched: 9793542 seqs, 4134689005 residues

Word size : 0
Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 50 summaries

Database : Published Applications NA Main:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	207	100.0	207	8	US-10-674-124A-2
2	50	24.2	1026	3	US-09-866-987-1
3	37	17.9	52138	10	US-11-166-990-75
4	36	17.4	67093	8	US-10-417-375-33
5	33	15.9	339234	7	US-10-322-696-73
6	32	15.5	447	8	US-10-674-124A-555
7	32	15.5	1742	4	US-09-925-065A-678678
8	32	15.5	39835	6	US-10-085-117-277
9	32	15.5	70019	7	US-10-322-281-823
10	32	15.5	77530	7	US-10-322-281-265
11	31	15.0	882	4	US-09-925-065A-72245
12	31	15.0	882	4	US-09-925-065A-72246
13	31	15.0	1475	4	US-09-925-065A-670284
14	31	15.0	1475	4	US-09-925-065A-670285
15	31	15.0	1475	4	US-09-925-065A-670286
16	30	14.5	402	3	US-09-983-965-5596
17	29	14.0	253	8	US-10-674-124A-6084
18	29	14.0	505	5	US-10-027-632-6710
19	29	14.0	505	6	US-10-027-632-6710
20	29	14.0	609	5	US-10-027-632-222642
21	29	14.0	609	6	US-10-027-632-222642
22	29	14.0	611	4	US-09-925-065A-588742
23	29	14.0	1194	5	US-10-027-632-119333

C 24	29	14.0	1194	5	US-10-027-632-119334	Sequence 119334,
C 25	29	14.0	1194	5	US-10-027-632-119335	Sequence 119335,
C 26	29	14.0	1194	6	US-10-027-632-119333	Sequence 119333,
C 27	29	14.0	1194	6	US-10-027-632-119334	Sequence 119334,
C 28	29	14.0	1194	6	US-10-027-632-119335	Sequence 119335,
C 29	29	14.0	1249	4	US-09-925-065A-667972	Sequence 667972,
C 30	29	14.0	1703	4	US-09-925-065A-76909	Sequence 76909, A
C 31	29	14.0	1724	4	US-09-925-065A-78157	Sequence 78157, A
C 32	29	14.0	1724	4	US-09-925-065A-78158	Sequence 78158, A
C 33	29	14.0	4361	10	US-11-097-143-5671	Sequence 5671, Ap
C 34	29	14.0	168050	9	US-10-939-321-1	Sequence 1, Appl
C 35	28	13.5	261	8	US-10-674-124A-16843	Sequence 16843, A
C 36	28	13.5	327	8	US-10-674-124A-13270	Sequence 13270, A
C 37	28	13.5	355	8	US-10-674-124A-4507	Sequence 4507, Ap
C 38	28	13.5	454	4	US-09-925-065A-464152	Sequence 464152,
C 39	28	13.5	464	4	US-09-925-065A-556697	Sequence 556697,
C 40	28	13.5	502	4	US-09-925-065A-835230	Sequence 835230,
C 41	28	13.5	517	4	US-09-925-065A-763473	Sequence 763473,
C 42	28	13.5	517	4	US-09-925-065A-763474	Sequence 763474,
C 43	28	13.5	760	4	US-09-925-065A-714119	Sequence 714119,
C 44	28	13.5	881	5	US-10-027-632-154764	Sequence 154764,
C 45	28	13.5	881	6	US-10-027-632-154764	Sequence 27553, A
C 46	28	13.5	22788	10	US-11-097-143-27553	Sequence 861, App
C 47	28	13.5	33916	7	US-10-322-281-861	Sequence 24, Appl
C 48	28	13.5	39699	7	US-10-322-281-24	Sequence 17854, A
C 49	28	13.5	49431	10	US-11-097-143-17854	Sequence 16, Appl
C 50	28	13.5	96593	3	US-09-997-722-16	

ALIGNMENTS

RESULT 1
US-10-674-124A-2
; Sequence 2, Application US/10674124A
; Publication No. US2004019797A1
; GENERAL INFORMATION:
; APPLICANT: INOKO, Hidetoshi
; APPLICANT: TAMIYA, Gen
; TITLE OF INVENTION: GENE MAPPING METHOD USING MICROSATELLITE
; FILE OF INVENTION: GENETIC POLYMORPHISM MARKERS
; FILE REFERENCE: ORIN-003CIP
; CURRENT APPLICATION NUMBER: US/10/674,124A
; CURRENT FILING DATE: 2003-09-26
; PRIOR APPLICATION NUMBER: 10/257,511
; PRIOR FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: PCT/JP00/07621
; PRIOR FILING DATE: 2000-10-30
; PRIOR APPLICATION NUMBER: JP2000-112699
; PRIOR FILING DATE: 2000-04-13
; PRIOR APPLICATION NUMBER: JP2002-327516
; PRIOR FILING DATE: 2002-09-28
; PRIOR APPLICATION NUMBER: JP2002-383869
; PRIOR FILING DATE: 2002-12-09
; NUMBER OF SEQ ID NOS: 27110
; SEQ ID NO 2
; LENGTH: 207
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Located on chromosome 1
; FEATURE:
; OTHER INFORMATION: Distance between a terminus base of telomere on
; OTHER INFORMATION: chromosomal short arm and 5'-terminus of this base
; OTHER INFORMATION: sequence : 100269
; FEATURE:
; OTHER INFORMATION: Distance between 3'-terminus of neighbour sequence of
; OTHER INFORMATION: sequence listing upward to telomere on chromosomal short arm and
; OTHER INFORMATION: 5'-terminus of this base sequence : 91592
US-10-674-124A-2

```
Query Match      100.0%; Score 207; DB 8; Length 207;
Best Local Similarity 100.0%; Pred. No. 1e-100;
Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AACCAAGTCTTAGCAGACGGCGTCCCTCAGGATGCGGCACACGCGCACACACACACACA 60
   |||||||
DB 1 AACCAAGTCTTAGCAGACGGCGTCCCTCAGGATGCGGCACACGCGCACACACACACA 60

QY 61 CACATGACACGCGCACACATGAACACACATGGAACACGCGGAACACACATGACACAGTA 120
   |||||||
DB 61 CACATGACACGCGCACACATGAACACACATGGAACACGCGGAACACACATGACACAGTA 120

QY 121 CACACGTGTGCTGACACGCGGATGCTGGCCAGGGCTCACCATCCAAACGGGGGCTCTG 180
   |||||||
DB 121 CACACGTGTGCTGACACGCGGATGCTGGCCAGGGCTCACCATCCAAACGGGGGCTCTG 180

QY 181 GCAGCCACGCGGTACACAGACTCAAT 207
   |||||||
DB 181 GCAGCCACGCGGTACACAGACTCAAT 207

RESULT 2
US-09-866-987-1/c
; Sequence 1, Application US/09866987
; Patent No. US20020090703A1
; GENERAL INFORMATION:
; APPLICANT: PLOWMAN, GREGORY D.
; APPLICANT: MARTINEZ, RICARDO
; APPLICANT: WHYTE, DAVID
; APPLICANT: MANNING, GERARD
; APPLICANT: SUDARSANAM, SUCHA
; APPLICANT: CAENEPEEL, SEAN
; APPLICANT: HILL, RON
; APPLICANT: FLANAGAN, PETER
; TITLE OF INVENTION: MAMMALIAN PROTEIN PHOSPHATASES
; FILE REFERENCE: 038602/1180
; CURRENT APPLICATION NUMBER: US/09/866,987
; CURRENT FILING DATE: 2001-05-30
; PRIOR FILING DATE: 60/208,291
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1026
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: phosphatase nucleotide
US-09-866-987-1

Query Match      24.2%; Score 50; DB 3; Length 1026;
Best Local Similarity 100.0%; Pred. No. 1.1e-16;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 158 CACCACTTCAACGGCGGGTCTTGGCAGCCCGGCTACCAAGCTCAAT 207
   |||||||
DB 323 CACCACTTCAACGGCGGGTCTTGGCAGCCCGGCTACCAAGCTCAAT 274

RESULT 3
US-11-166-990-75
; Sequence 75, Application US/11166990
; Publication No. US20050239125A1
; GENERAL INFORMATION:
; APPLICANT: Hodges, Timothy A.
; TITLE OF INVENTION: METHODS FOR GENOTYPE SCREENING
; FILE REFERENCE: 023131.41500
; CURRENT APPLICATION NUMBER: US/11/166,990
; CURRENT FILING DATE: 2005-06-24
; PRIOR APPLICATION NUMBER: 60/230,371
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 09/945,952
```

```
Query Match      17.9%; Score 37; DB 10; Length 52138;
Best Local Similarity 100.0%; Pred. No. 5.6e-10;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 45 CACACACACACACACACATGCGACGCGCACACACAT 81
   |||||||
DB 16029 CACACACACACACACACATGCGACGCGCACACACAT 16065

RESULT 4
US-10-417-375-33/c
; Sequence 33, Application US/10417375
; Publication No. US20040219528A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc Malandro
; TITLE OF INVENTION: Novel Therapeutic Targets in Cancer
; FILE REFERENCE: 529452001600
; CURRENT APPLICATION NUMBER: US/10/417,375
; CURRENT FILING DATE: 2003-04-15
; NUMBER OF SEQ ID NOS: 176
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 33
; LENGTH: 67093
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(67093)
; OTHER INFORMATION: n = A,T,C or G
US-10-417-375-33

Query Match      17.4%; Score 36; DB 8; Length 67093;
Best Local Similarity 100.0%; Pred. No. 1.9e-09;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 45 CACACACACACACACACATGCGACGCGCACACA 80
   |||||||
DB 29466 CACACACACACACACACATGCGACGCGCACACA 29431

RESULT 5
US-10-322-696-73/c
; Sequence 73, Application US/10322696
; Publication No. US20040166490A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Malandro, Marc
; TITLE OF INVENTION: NOVEL THERAPEUTIC TARGETS IN CANCER
; FILE REFERENCE: 529452001200
; CURRENT APPLICATION NUMBER: US/10/322,696
; CURRENT FILING DATE: 2003-10-17
; NUMBER OF SEQ ID NOS: 186
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 73
; LENGTH: 339234
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

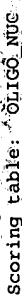
OM nucleic - nucleic search, using sw model

Run on: December 14, 2005, 03:22:04 ; Search time 67.3596 Seconds
(without alignments)
5462.552 Million cell updates/sec

Title: US-10-674-124A-2

Perfect score: 207

Sequence: 1 aaccagtgcttagcagacgc.....agcgggtaccagagctcaat 207

Scoring table: 

Searched: 1303057 seqs, 888780828 residues

Word size: 6, 0

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 50 summaries

Database : Issued Patents NA.*

- 1: /cgn2_6/ptodata/1/ina/1 COMB.seq.*
- 2: /cgn2_6/ptodata/1/ina/5 COMB.seq.*
- 3: /cgn2_6/ptodata/1/ina/6A COMB.seq.*
- 4: /cgn2_6/ptodata/1/ina/6B COMB.seq.*
- 5: /cgn2_6/ptodata/1/ina/6H COMB.seq.*
- 6: /cgn2_6/ptodata/1/ina/PP COMB.seq.*
- 7: /cgn2_6/ptodata/1/ina/RE COMB.seq.*
- 8: /cgn2_6/ptodata/1/ina/RE COMB.seq.*
- 9: /cgn2_6/ptodata/1/ina/backfiles.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	33	15.9	268	3	US-09-513-999C-8748, Ap
C 2	32	15.5	601	3	US-09-949-016-132718
C 3	32	15.5	601	3	US-09-949-016-132719
C 4	32	15.5	601	3	US-09-949-016-132720
C 5	32	15.5	194714	3	US-09-949-016-11869
C 6	32	15.5	196714	3	US-09-949-016-15474
C 7	30	14.5	601	3	US-09-949-016-109508
C 8	30	14.5	601	3	US-09-949-016-126979
C 9	30	14.5	601	3	US-09-949-016-136980
C 10	30	14.5	32202	3	US-09-949-016-15357
C 11	30	14.5	67581	3	US-09-949-016-14768
C 12	29	14.5	67843	3	US-09-949-016-14760
C 13	29	14.0	175236	3	US-09-949-016-14353
C 14	28	13.5	601	3	US-09-949-016-62776
C 15	28	13.5	601	3	US-09-949-016-62777
C 16	28	13.5	601	3	US-09-949-016-62778
C 17	28	13.5	601	3	US-09-949-016-62779
C 18	28	13.5	601	3	US-09-949-016-192212
C 19	28	13.5	601	3	US-09-949-016-192213
C 20	28	13.5	13948	3	US-09-949-016-13023
C 21	28	13.5	90776	3	US-09-949-016-17230
C 22	28	13.5	103792	3	US-09-949-016-13553
C 23	28	13.5	103993	3	US-09-949-016-12317
C 24	27	13.0	44	3	US-09-443-199C-1017

Sequence 1017, Ap
Sequence 23415, A
Sequence 257, App
Sequence 385, App
Sequence 715, App
Sequence 14774, A
Sequence 29, Appl
Sequence 6074, Ap
Sequence 21356, A
Sequence 336, App
Sequence 58095, A
Sequence 58096, A
Sequence 141439,
Sequence 142850,
Sequence 142851,
Sequence 152327,
Sequence 162901,
Sequence 162902,
Sequence 182840,
Sequence 198491,
Sequence 198492,
Sequence 198491,
Sequence 2, Appli
Sequence 12075, A
Sequence 14226, A
Sequence 208, App
Sequence 208, App

ALIGNMENTS

RESULT 1

US-09-513-999C-8748/c

; Sequence 8748, Application US/09513999C

; Patent No. 6783361

; GENERAL INFORMATION:

; APPLICANT: Dumas Milne Edwards, J.B.

; APPLICANT: Duclert, A.

; APPLICANT: Giordano, J.Y.

; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.

; Patent No. 6783361

; FILE REFERENCE: 59. US2.REG

; CURRENT APPLICATION NUMBER: US/09/513.999C

; CURRENT FILING DATE: 2000-02-24

; PRIOR APPLICATION NUMBER: US 60/122,487

; PRIOR FILING DATE: 1999-02-26

; NUMBER OF SEQ ID NOS: 36681

; SOFTWARE: Patent.pm

; SEQ ID NO 8748

; LENGTH: 268

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc_feature

; LOCATION: 168

; OTHER INFORMATION: k=g or t

US-09-513-999C-8748

Query Match 15.9%; Score 33; DB 3; Length 268;

Best Local Similarity 100.0%; Pred. No. 1.5e-06;

Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 ACACACACACACATGCACACGACACATG 82

Db 71 ACACACACACACATGCACACGACACATG 39

RESULT 2

US-09-949-016-132718

; Sequence 132718, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

```
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 132718
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-132718

Query Match      15.5%; Score 32; DB 3; Length 601;
Best Local Similarity 100.0%; Pred. No. 4.3e-06;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      45 CACACACACACACACATGCACACGCACA 76
Db      47 CACACACACACACACATGCACACGCACA 78

RESULT 3
US-09-949-016-132719
; Sequence 132719, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 132719
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-132719

Query Match      15.5%; Score 32; DB 3; Length 601;
Best Local Similarity 100.0%; Pred. No. 4.3e-06;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      45 CACACACACACACACATGCACACGCACA 76
Db      63 CACACACACACACACATGCACACGCACA 94

RESULT 4
US-09-949-016-132720
; Sequence 132720, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
```

```
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 132720
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-132720

Query Match      15.5%; Score 32; DB 3; Length 601;
Best Local Similarity 100.0%; Pred. No. 4.3e-06;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      45 CACACACACACACACATGCACACGCACA 76
Db      208 CACACACACACACACATGCACACGCACA 239

RESULT 5
US-09-949-016-11869/c
; Sequence 11869, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11869
; LENGTH: 194714
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-11869

Query Match      15.5%; Score 32; DB 3; Length 194714;
Best Local Similarity 100.0%; Pred. No. 2.5e-06;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      45 CACACACACACACACATGCACACGCACA 76
Db      15608 CACACACACACACACATGCACACGCACA 15577

RESULT 6
US-09-949-016-15474/c
; Sequence 15474, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 14, 2005, 04:20:34 ; Search time 1927.59 Seconds
(without alignments)
5024.367 Million cell updates/sec

Title: US-10-674-124A-2
Perfect score: 207
Sequence: 1 aaccagtgttagcagcgc.....agcgggtaccagactcaat 207

Scoring table: OligoNUC
Gapop 60.0, Gapext 60.0

Searched: 41078325 seqs, 23393541228 residues

Word size: 20

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 50 summaries

Database : EST: *
1: gb_est1.*
2: gb_est2.*
3: gb_est3.*
4: gb_hic.*
5: gb_est4.*
6: gb_est5.*
7: gb_est6.*
8: gb_est7.*
9: gb_ges1.*
10: gb_ges2.*
11: gb_ges3.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	38	18.4	671	9	AZ587362
2	36	17.4	287	9	AZ733430
3	36	17.4	478	3	BO194283
4	36	17.4	627	11	CNS03DHD
5	36	17.4	840	11	CR169879
6	36	17.4	1042	11	CNS04CZG
7	36	17.4	1049	11	CNS05DIS
8	35	16.9	283	10	BX972640
9	35	16.9	494	9	AZ407953
10	35	16.9	496	9	AZ024543
11	35	16.9	611	9	AZ380969
12	35	16.9	1201	10	AG338469
13	34	16.4	376	3	BI534913
14	34	16.4	415	6	CB799941
15	34	16.4	486	10	BX177437
16	34	16.4	502	9	BZ148403
17	34	16.4	510	11	FR0043263
18	34	16.4	537	9	AZ006109
19	34	16.4	572	11	FR0043899
20	34	16.4	600	11	FR0041672
21	34	16.4	618	11	FR0043913
22	34	16.4	713	9	BZ282696

C 23	34	16.4	754	11	CR128435
C 24	34	16.4	769	9	BZ139452
C 25	34	16.4	792	9	BZ282602
C 26	34	16.4	804	11	CR190843
C 27	34	16.4	901	7	CK401939
C 28	33	15.9	307	9	AZ022678
C 29	33	15.9	475	9	AQ285135
C 30	33	15.9	523	9	AQ148705
C 31	33	15.9	622	9	AZ004814
C 32	33	15.9	629	9	AQ194800
C 33	33	15.9	700	10	AG341713
C 34	33	15.9	739	10	AG516767
C 35	33	15.9	888	6	CA511475
C 36	33	15.9	1101	10	CNS014NB
C 37	32	15.5	221	9	AZ436300
C 38	32	15.5	291	2	BB526865
C 39	32	15.5	292	2	BB511252
C 40	32	15.5	378	2	BE849531
C 41	32	15.5	403	1	AI037508
C 42	32	15.5	407	1	AI664473
C 43	32	15.5	424	9	BZ251046
C 44	32	15.5	469	1	AW989187
C 45	32	15.5	473	1	AW913500
C 46	32	15.5	473	2	BE688977
C 47	32	15.5	498	1	AW990071
C 48	32	15.5	547	5	BX517541
C 49	32	15.5	547	6	CD775323
C 50	32	15.5	549	1	AI316720

ALIGNMENTS

AZ587362 671 bp DNA linear GSS 13-DEC-2000
1M0395C05F Mouse 10kb plasmid UUGCLM library Mus musculus genomic
clone UUGC1M0395C05 F, genomic survey sequence.

AZ587362
GSS.
Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Mus.

1 (bases 1 to 671)
Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,
Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T.,
Reilly, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von
Niederhausern, A., and Wright, D., Weiss, R.

Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
Unpublished (2000)

JOURNAL

Contact: Robert B. Weiss

University of Utah Genome Center

University of Utah

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT

84112, USA

Tel: 801 585 5606

Fax: 801 585 7177

Email: ddunn@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00

Plate: 0395 row: C column: 05

Seq primer: CATTGTAAACGACGCCAGT

Class: plasmid ends

High quality sequence stop: 671.

Location/Qualifiers

1. .671

/organism="Mus musculus"

/mol_type="genomic DNA"

/strain="C57BL/6J"

/db_xref="taxon:10090"

```

/clone="UUGC1M0395C05"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/notes="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PWD42 [gi|4732114|gb|AF129072.1], a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

ORIGIN
Query Match 18.4%; Score 38; DB 9; Length 671;
Best Local Similarity 100.0%; Pred. No. 3.3e-08;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 45 CACACACACACACACATGCACGACACACATG 82
|||||
Db 282 CACACACACACACACATGCACGACACACATG 319

RESULT 2
LOCUS AZ733430 287 bp DNA linear GSS 25-JAN-2001
DEFINITION RPCI-24-118H15-TV RPCI-24 Mus musculus genomic clone
ACCESSION AZ733430
VERSION AZ733430.1 GI:12497660
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Mus.
1 (bases 1 to 287)
Zhao, S., Nierman, W., Malek, J., Shatsman, S., Akinret, B., Levins, M.,
Tsegaye, G., Geer, K., Krol, M., Shvartsbeyn, A., Gebregeorgis, E.,
Russell, D., de Jong, P. and Fraser, C.M.
Mouse BAC End Sequences from Library RPCI-24
Unpublished (1999)
Other GSSs: RPCI-24-118H15.TJ
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: shao@tigr.org
Clones are derived from the mouse BAC library RPCI-24. For BAC
library availability, please contact Pieter de Jong
(pdejong@mail.cho.org). Clones may be purchased from BACPAC
Resources (http://www.choi.org/bacpac/orderingframe.htm). BAC end
page: http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
Plate: 118 row: H column: 15
Seq primer: T7
Class: BAC ends.
Location/Qualifiers
1. .287
/organism="Mus musculus"
/mol_type="genomic DNA"

/clone="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPCI-24-118H15"
/sex="Male"
/cell_type="Spleen/Brain"
/clone_lib="RPCI-24"
/notes="Vector: pTARBAC1; Site 1: BamH1; Site 2: BamH1;
RPCI-24 Mouse BAC library produced by Pieter de Jong. The
library was cloned in the pTARBAC1 cloning vector at the
BamH1 sites using MboI partially digested male C57BL/6J
DNA."

ORIGIN
Query Match 17.4%; Score 36; DB 9; Length 287;
Best Local Similarity 100.0%; Pred. No. 3.5e-07;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 45 CACACACACACACACATGCACGACACACACA 80
|||||
Db 234 CACACACACACACACATGCACGACACACACA 199

RESULT 3
LOCUS BQ194283 478 bp mRNA linear EST 30-APR-2002
DEFINITION UI-R-CNI-cma-e-19-0-UI.s1 UI-R-CNI Rattus norvegicus cDNA clone
ACCESSION BQ194283
VERSION BQ194283.1 GI:20369834
KEYWORDS EST.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
1 (bases 1 to 478)
Bonaldo, M.F., Lennon, G. and Soares, M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
8889548
Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
The sequence contained an oligo-dT track that was present in the
oligonucleotide that was used to prime the synthesis of first
strand cDNA and therefore this may represent a bonafide poly A
tail. The sequence tag present in the cDNA between the NotI site
and the oligo-dT track served to identify it as a clone from the
normalized fundus library cDNA Library Preparation: M.B. Soares Lab
Clone distribution: clones will be available through Research
Genetics (www.resgen.com) The following repetitive elements were
found in this cDNA sequence: 14-84, >(CA)n#Simple_repeat 86-356,
>B4A#SINE/B4
Seq primer: M13 Forward
POLYA=Yes.
Location/Qualifiers
1. .478
/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UI-R-CNI-cma-e-19-0-UI"
/dev_stage="adult"
/lab_host="DHI0B (Life Technologies)"
/clone_lib="UI-R-CNI"
/notes="Vector: pT7T3D-pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; The UI-R-CNI
library is a subtracted library derived from the following

```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model
Run on: December 14, 2005, 03:13:04 ; Search time 228.838 Seconds
(without alignments)
6028.680 Million cell updates/sec

Title: US-10-674-124A-2
Perfect score: 207
Sequence: 1 aacagtgcttagacagc.....agcgggtaccagagctcaat 207

Scoring table: OMF000-NUC
Gapop 60 0 0 , Gapext 60.0

Searched: 4996997 seqs, 332346308 residues

Word size: 3 0 0

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 50 summaries

- Database : N Geneseq 21:*
- 1: geneseqn1980s:*
 - 2: geneseqn1990s:*
 - 3: geneseqn2000s:*
 - 4: geneseqn2001as:*
 - 5: geneseqn2001bs:*
 - 6: geneseqn2002as:*
 - 7: geneseqn2002bs:*
 - 8: geneseqn2003as:*
 - 9: geneseqn2003bs:*
 - 10: geneseqn2003cs:*
 - 11: geneseqn2003ds:*
 - 12: geneseqn2004as:*
 - 13: geneseqn2004bs:*
 - 14: geneseqn2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	50	24.2	1026	6 ABA94424	Abd32671 Human pro
C 2	36	17.4	67093	13 ABD32671	Abd32671 Mouse can
C 3	33	15.9	268	3 AAC04673	AAC04673 Human sec
C 4	33	15.9	90220	6 ABK83576	ABK83576 Human cDN
C 5	33	15.9	339234	12 ADQ59437	Adg59437 Human can
C 6	33	15.9	339234	14 AD213744	Ad213744 Murine ca
C 7	32	15.5	3267	12 ADK52131	Adk52131 Mouse ato
C 8	32	15.5	70019	13 ABD33601	Abd33601 Murine ca
C 9	32	15.5	77530	13 ABD33240	Abd33240 Murine ca
C 10	31	15.0	972	10 ADC12701	Adc12701 Human GPC
C 11	31	15.0	77287	9 AAD58261	Aad58261 Murine tu
C 12	31	15.0	167739	9 AAD58258	Aad58258 Bovine ES
C 13	30	14.5	402	8 ABX55667	Abx55667 Mouse can
C 14	30	14.5	38920	12 ADQ97252	Adq97252 Human pol
C 15	29	14.0	407	4 AAI90660	Aai90660 Human pol
C 16	29	14.0	4361	4 ABL05620	AbL05620 Drosophil
C 17	29	14.0	110000	12 ADQ97047	Adg97047 Mouse can
C 18	29	14.0	151052	12 ADQ97526	Adg97526 Mouse can
C 19	28	13.5	694	4 AAI86049	Aai86049 Human pol

C 20	28	13.5	22788	4 ABL20208	Abl20208 Drosophil
C 21	28	13.5	33916	13 ABD33626	Abd33626 Murine ca
C 22	28	13.5	33916	14 AD213599	Ad213599 Murine ca
C 23	28	13.5	39699	13 ABD33084	Abd33084 Human can
C 24	28	13.5	49431	4 ABD13742	AbL13742 Drosophil
C 25	28	13.5	96593	9 ADA02885	Ada02885 Mouse Blm
C 26	28	13.5	96593	9 ADA02750	Ada02750 Human IRF
C 27	28	13.5	96593	10 ADB72488	Adb72488 Human IRF
C 28	28	13.5	96593	10 ADB72623	Adb72623 Mouse Blm
C 29	28	13.5	96593	10 ADC85230	Adc85230 Human IRF
C 30	28	13.5	96593	12 ADM74345	Adm74345 Human car
C 31	28	13.5	96593	12 ADM74480	Adm74480 Murine ca
C 32	28	13.5	96594	10 ADC85364	Adc85364 Human Pap
C 33	28	13.5	110000	12 ADQ59443	AdQ59443 1
C 34	28	13.5	110000	14 AD213754	Ad213754 1
C 35	28	13.5	113575	11 ACN44348	AcN44348 Mouse gen
C 36	28	13.5	322885	13 ADS93537	AdS93537 Human MRC
C 37	27	13.0	44	3 AAA77336	Aaa77336 Human clo
C 38	27	13.0	48	3 AAA77334	Aaa77334 Human cDN
C 39	27	13.0	51	4 AAL27990	Aal27990 Human SNP
C 40	27	13.0	170	3 AAC19340	Aac19340 Human sec
C 41	27	13.0	279	3 AAA74225	Aaa74225 Lobiolly
C 42	27	13.0	286	3 AAC00387	Aac00387 Human sec
C 43	27	13.0	299	2 AAV87891	Aav87891 EST clone
C 44	27	13.0	310	3 AAZ80631	Aaz80631 Human col
C 45	27	13.0	314	3 AAC10699	Aac10699 Human sec
C 46	27	13.0	333	10 AD443552	Ad443552 Mouse PAP
C 47	27	13.0	333	13 ADU47661	Adu47661 Mouse chr
C 48	27	13.0	347	4 AAI81990	Aai81990 Human pol
C 49	27	13.0	358	14 ADW82173	Adw82173 MAP3K9 ma
C 50	27	13.0	381	4 AAI86122	Aai86122 Human pol

ALIGNMENTS

RESULT 1

ABA94424/c

ID ABA94424 standard; cDNA; 1026 BP.

XX ABA94424;

XX 26-MAR-2002 (first entry)

XX Human protein phosphatase SGP057 encoding cDNA.

XX Protein phosphatase; protein tyrosine phosphatase; PTP; STP; SGP057; cytosolic; immunosuppressive; vasotropic; analgesic; hypertensive; KW hypotensive; nootropic; neuroprotective; virucide; antibacterial; human; KW fungicide; ophthalmological; antiinflammatory; cardiac; antimigraine; KW antiarthritic; muscular; antiasthmatic; osteopathic; antipsoriatic; KW antiarteriosclerotic; antiallergic; antidiabetic; tranquilizer; cancer; KW neuroleptic; antiparkinsonian; anorectic; gene therapy; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

CDS 1..1026

/*tag= a

/product= "SGP057"

XX WO200192491-A2.

XX 06-DEC-2001.

XX 30-MAY-2001; 2001WO-US017305.

XX 30-MAY-2000; 2000US-0208291P.

XX (SUGS-) SUGEN INC.

XX Plowman GD, Whyte D, Sudarsanam S, Manning G, Caenepel S;

XX Martinez R, Hill RJ, Planagan P;

```
XX WPI; 2002-114340/15.
DR P-FSDB; ABB07302.
XX
XX New phosphatase polypeptides for treating e.g., cancers, immune-related
PT diseases and disorders, cardiovascular diseases, or brain or neuronal-
PT associated diseases.
XX
XX Claim 28; Fig 1A; 161pp; English.
XX
XX The invention relates to human protein phosphatase polypeptides and
CC polynucleotides. The phosphatase polypeptides are useful for diagnosing
CC and treating various phosphatase-related diseases and conditions, such as
CC cancers, immune-related diseases and disorders, cardiovascular diseases
CC (e.g., reperfusion restenosis), brain or neuronal-associated diseases,
CC metabolic disorders (e.g., diabetes or obesity), central or peripheral
CC nervous system diseases and conditions (e.g., pain, migraine, mood
CC disorders, hypotension, or hypertension), psychotic and neurological
CC disorders (e.g., anxiety, schizophrenia), neurodegenerative diseases
CC (e.g., Alzheimer's or Parkinson's disease), infections caused by viral
CC (e.g., human immunodeficiency virus (HIV)), bacterial, fungal or prion-
CC agents, ocular diseases, macular degeneration, and inflammatory disorders
CC (e.g., rheumatoid arthritis, multiple sclerosis, osteoarthritis). The
CC polypeptides may also be used for screening for small molecule compounds
CC that inhibit the catalytic activity of the encoded protein phosphatase
CC with potential use for treating the above-mentioned diseases, for
CC generating antibodies or hybridomas, and as a diagnostic tool for
CC detecting a disease or disorder. Genetic sequences of phosphatases are
CC useful in gene therapy. The present sequence represents a cDNA encoding a
CC novel human protein tyrosine phosphatase (PTP)-like phosphatase
CC polypeptide SGP057
XX
XX Sequence 1026 BP; 176 A; 294 C; 393 G; 163 T; 0 U; 0 Other;
SQ
Query Match 24.2%; Score 50; DB 6; Length 1026;
Best Local Similarity 100.0%; Pred. No. 9.3e-15;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 158 CACCACCTCCACGGGGTCTCTGGCAGCCCGGGTACCGAGCTCAAT 207
DB 323 CACCACCTCCACGGGGTCTCTGGCAGCCCGGGTACCGAGCTCAAT 274
RESULT 2
ABD32671/c
ID ABD32671 standard; DNA; 67093 BP.
XX
XX ABD32671;
XX
XX 18-NOV-2004 (first entry)
XX
XX Mouse cancer-associated genomic DNA MD13-075.
XX
XX Mouse; ds; cancer-associated protein; gene; cytostatic; cancer;
XX leukaemia; lymphoma; CAP.
XX
XX Mus musculus.
XX
XX WO2004074320-A2.
XX
XX 02-SEP-2004.
XX
XX 17-FEB-2004; 2004WO-US004730.
XX
XX 14-FEB-2003; 2003US-00367094.
XX 14-MAR-2003; 2003US-00388838.
XX 15-APR-2003; 2003US-00417375.
XX 13-JUN-2003; 2003US-00461862.
XX 15-SEP-2003; 2003US-00663431.
XX 15-DEC-2003; 2003US-00737318.
XX (SAGR-) SAGRES DISCOVERY INC.
XX
XX
XX Morris DW, Morris DW, Malandro MS;
XX WPI; 2004-652914/63.
XX
XX New isolated cancer-associated polynucleotides and polypeptides useful
PT for diagnosing, preventing or treating cancers, especially lymphoma and
PT leukemia, or in screening for agents that modulate cancer.
XX
XX disclosure; seqid 218; 310pp; English.
XX
XX The invention relates to an isolated nucleic acid comprising at least 10
CC contiguous nucleotides of any of the 233 polynucleotide sequences given
CC in the specification, or its complement. The nucleic acids encode cancer-
CC associated proteins. Also included are an expression vector comprising
CC the isolated nucleic acid cited above, a host cell comprising the above
CC recombinant nucleic acid or expression vector, a microarray for detecting
CC a cancer-associated (CA) nucleic acid comprising at least one probe
CC comprising at least 10 contiguous nucleotides of any of the above-
CC mentioned nucleotide sequences, an isolated polypeptide (encoded within
CC an open reading frame of a CA sequence selected from any of the 95
CC polynucleotide sequences as mentioned in the specification, or its
CC complement), an isolated antibody, (or its antigen binding fragment) that
CC binds to the above polypeptide, a hybridoma that produces the above
CC monoclonal antibody, a pharmaceutical composition comprising the above
CC antibody and a pharmaceutical excipient, a kit for detecting cancer
CC cells (comprising the antibody cited above, methods for diagnosing cancer
CC or for detecting the presence or absence of cancer cells in an
CC individual, a method for inhibiting growth of cancer cells in an
CC individual, a method for delivering a therapeutic agent to cancer cells
CC in an individual, an electronic library comprising the above
CC polynucleotide or polypeptide (or their fragments), methods of screening
CC for anticancer activity or for a bioactive agent capable of modulating
CC the activity of a CA protein (CAP), methods for detecting cancer
CC associated with expression of a polypeptide in a test cell sample, a
CC method for treating cancers and a method for inhibiting the expression of
CC CA gene in a cell. The composition and methods are useful for detecting,
CC diagnosing, preventing and treating cancers, especially lymphoma and
CC leukaemia. These may also be used in screening for agents that modulate
CC cancer. The present sequence is a mouse CAP genomic sequence. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 67093 BP; 17235 A; 15136 C; 16131 G; 18531 T; 0 U; 60 Other;
SQ
Query Match 17.4%; Score 36; DB 13; Length 67093;
Best Local Similarity 100.0%; Pred. No. 6.4e-08;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 45 CACACACACACACACACATGCACACGCGACACACA 80
DB 29466 CACACACACACACACACATGCACACGCGACACACA 29431
RESULT 3
AAC04673/c
ID AAC04673 standard; cDNA; 268 BP.
XX
XX AAC04673;
XX
XX 06-OCT-2000 (first entry)
XX
XX Human secreted protein 5' EST, SEQ ID NO: 8748.
XX
XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
XX gene therapy; chromosome mapping; ss.
XX
XX Homo sapiens.
XX
XX EP1033401-A2.
XX
XX 06-SEP-2000.
XX
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 13, 2005, 22:13:37 ; Search time 1202.94 Seconds
(without alignments)
9781.550 Million cell updates/sec

Title: US-10-674-124A-2
Perfect score: 207
Sequence: 1 aaccagtgcttagcagacgc.....agcgggtaccagagctcaat 207

Scoring table: OLIGO NUC
Gapop 60.0 , Gapext 60.0

Searched: 5883141 seqs, 28421725653 residues

Word size : 0
Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 50 summaries

Database : GenEmbl.*
1: gb_ba.*
2: gb_in.*
3: gb_env.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pr.*
9: gb_ro.*
10: gb_sts.*
11: gb_sv.*
12: gb_un.*
13: gb_vi.*
14: gb_htg.*
15: gb_pl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	207	100.0	207	10	AB151030
C 2	207	100.0	130979	8-	AL3390719
C 3	161	77.8	327883	14	BX571684
C 4	50	24.2	1026	6	AX352512
C 5	44	21.3	224653	14	AC149246
C 6	43	20.8	159031	14	CR339053
C 7	39	18.8	199235	9	AC110164
C 8	38	18.4	96128	9	AL663077
C 9	38	18.4	142364	9	AC156951
C 10	38	18.4	149347	14	AC131813
C 11	38	18.4	168635	14	AC142061
C 12	38	18.4	174894	14	AC118979
C 13	38	18.4	212022	14	AC134726
C 14	38	18.4	217297	14	AC114141
C 15	38	18.4	225524	9	AC133489
C 16	38	18.4	226183	14	AC123101
C 17	38	18.4	244692	14	AC128181
C 18	38	18.4	244707	14	AC098994

19	17.9	150109	9	AC132334
20	17.9	175275	14	AC091883
21	17.9	178488	9	AC123548
22	17.9	181955	8	AC010546
23	17.9	203193	14	AC128175
24	17.9	203193	14	AC128175
25	17.9	271477	14	AC121616
26	17.4	258	10	AU048719
27	17.4	446	10	G46315
28	17.4	67093	6	Q0869797
29	17.4	107113	14	AC142559
30	17.4	110000	14	AC110125_1
31	17.4	110000	14	AC128511_2
32	17.4	110000	14	AC130034_1
33	17.4	115003	5	EX005089
34	17.4	120487	9	AC103637
35	17.4	131042	9	AC110166
36	17.4	135686	14	AC140742
37	17.4	135800	14	AF523316
38	17.4	143147	14	AC141204
39	17.4	146823	14	AC161294
40	17.4	147370	9	AC123039
41	17.4	148976	14	AC130201
42	17.4	153492	9	AC124391
43	17.4	154614	14	CR855263
44	17.4	155547	9	AC132595
45	17.4	157393	9	AC124199
46	17.4	160832	14	AC142051
47	17.4	162572	14	AC166627
48	17.4	164485	9	AC122513
49	17.4	165196	9	AL596123
50	17.4	166292	9	AC123049
	17.4	169501	9	AC125100

ALIGNMENTS

RESULT 1

AB151030/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

AB151030 Homo sapiens DNA, STS on chromosome 1, D1S14251, sequence tagged site.
207 bp DNA linear STS 02-APR-2005

AB151030 GI:62171848

AB151030.1

STS.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.

1

Tamiya, G., Shinya, M., Ikuta, T., Makino, S., Matsumoto, T., Mano, S., Furuguchi, K., Ando, S., Nozaki, Y., Yukawa, W., Nakashige, R., Yamaguchi, D., Ishibashi, H., Yonekura, M., Nakami, Y., Takayama, S., Endo, T., Saruwatari, T., Yagura, M., Yoshikawa, Y., Fujimoto, K., Oka, A., Chiku, S., Linsen, S. E., Giphart, M. J., Bahram, S., Kulski, Y. J., Fukazawa, T., Hashimoto, H., Hoshina, Y., Suzuki, Y., Hotta, T., Mochida, J., Minezaki, T., Komai, K., Shiozawa, S., Taniguchi, A., Yamanaka, H., Kamatani, N., Imanishi, T., Gojobori, T. and Inoko, H.

A Whole Genome Association Study of Rheumatoid Arthritis using 27,039 Microsatellite Markers

Unpublished

2 (bases 1 to 207)

Tamiya, G., Makino, S., Fujimoto, K., Oka, A., Hayashi, H., Denda, A., Linsen, S. E., Ikuta, T., Shinya, M., Endo, T., Tomizawa, M., Tokubo, E., Sato, R., Takaki, A., Nagatsuka, Y., Watanabe, H., Adachi, S., Makino, Y., Nakano, S., Yamamoto, A., Yoshida, K., Okamoto, K., Yamaguchi, D., Ishibashi, H., Yonekura, M., Takayama, S., Nakami, Y., Saruwatari, T., Brand, A., van Hilten, J. A., van de Watering, L. M., Giphart, M. J., Bahram, S., Kulski, Y. J. and Inoko, H.

Direct Submission

Submitted (12-DEC-2003) Hidetoshi Inoko, Tokai University School of

Medicine, Department of Genetic Information; Bohseidai, Isehara, Kanagawa, 259-1193, Japan (E-mail: hinoko@is.icc.u-tokai.ac.jp, Tel: 81-463-93-1121, Fax: 81-463-94-8884)

COMMENT

Polymorphisms were confirmed by comparing with the pooled DNA typing data of 88 Dutch population or 100 Australian. When there was no significant difference between the Japanese and these Caucasians, individual typing was performed to confirm polymorphisms.

FEATURES

source
1. .207
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="1"
/map="lp36.33"
/note="pooled DNA of 100 Japanese unrelated individuals
sequence tagged site D1S14251"
1. 18
/note="5' primer: ATGTAGCTCTGGTACCG"
/PCR_conditions="denaturation 96degC 5 min, 57degC 1 min, 72degC 1 min; 40 cycles 96degC 45 sec, 57degC 45 sec, 72degC 1 min"
144. .163
/note="sequence tagged site D1S14251"
/rpt_type=tandem
/rpt_unit="cg"
complement(189. .207)
/note="3' primer: AACCACTGCTTAGCAGACG"
primer_bind
1. 18
/note="5' primer: ATGTAGCTCTGGTACCG"
/PCR_conditions="denaturation 96degC 5 min, 57degC 1 min, 72degC 1 min; 40 cycles 96degC 45 sec, 57degC 45 sec, 72degC 1 min"
144. .163
/note="sequence tagged site D1S14251"
/rpt_type=tandem
/rpt_unit="cg"
complement(189. .207)
/note="3' primer: AACCACTGCTTAGCAGACG"
repeat_region
144. .163
/note="sequence tagged site D1S14251"
/rpt_type=tandem
/rpt_unit="cg"
complement(189. .207)
/note="3' primer: AACCACTGCTTAGCAGACG"
primer_bind
1. 18
/note="5' primer: ATGTAGCTCTGGTACCG"
/PCR_conditions="denaturation 96degC 5 min, 57degC 1 min, 72degC 1 min; 40 cycles 96degC 45 sec, 57degC 45 sec, 72degC 1 min"
144. .163
/note="sequence tagged site D1S14251"
/rpt_type=tandem
/rpt_unit="cg"
complement(189. .207)
/note="3' primer: AACCACTGCTTAGCAGACG"

ORIGIN

Query Match 100.0%; Score 207; DB 10; Length 207;
Best Local Similarity 100.0%; Pred. No. 1.7e-99;
Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AACCACTGCTTAGCAGACGCGCTCCCTCAGATGCGGCACACGACACACACACACACA 60
Db 207 AACCACTGCTTAGCAGACGCGCTCCCTCAGATGCGGCACACGACACACACACACA 148
Qy 61 CACATGACACGCGACACACATGACACATGACACGCGAACACACACATGACACGTA 120
Db 147 CACATGACACGCGACACACATGACACATGACACGCGAACACACACATGACACGTA 88
Qy 121 CACAGTGTGTGACACGCGATGCTGGCGAGGGCTCACACTCCAAACGGGGTCTG 180
Db 87 CACAGTGTGTGACACGCGATGCTGGCGAGGGCTCACACTCCAAACGGGGTCTG 28
Qy 181 GCAGCCGACGCGGTACACAGCTCAAT 207
Db 27 GCAGCCGACGCGGTACACAGCTCAAT 1

RESULT 2

AL390719/c
LOCUS
DEFINITION Human DNA sequence from clone RP11-465B22 on chromosome 1 Contains a novel gene, a novel gene, a protein tyrosine phosphatase family pseudogene, a novel gene (FLJ20584), a novel gene, a novel gene, a novel gene (FLJ36119) and five CpG islands, complete sequence.
ACCESSION
VERSION AL390719.47 GI:31071457
KEYWORDS HTG; FLJ20584; FLJ36119.
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 130979)
AUTHORS Matthews,L.
TITLE Direct Submission
JOURNAL Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vegas@sanger.ac.uk
Clone requests: clonerequests@sanger.ac.uk
On May 23, 2003 this sequence version replaced gi:30722520.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases:
Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr1
RP11-465B22 is from the library RP11-11.2 constructed by the group of Pieter de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pBACE3.6

----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: vegas@sanger.ac.uk

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

FEATURES

source

1. .130979
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="1"
/clone="RP11-465B22"
/clone_lib="RP11-11.2"
misc_feature
2000
/note="Clone right end: RP11-5407"
gene
join(AL645608.30:186058..186658,2152..2379)
/locus_tag="RP11-465B22.2-001"
mRNA
join(AL645608.30:186058..186658,2152..2379)
/locus_tag="RP11-465B22.2-001"
/product="putative novel transcript"
/notes="match: ESTs: Em:BQ918995.1"
gene
join(2167..2290,2571..2857,3301..3492,3581..3680,4443..4566,5678..5758,6657..6737,6960..7094,7668..7815,8299..8446)
/locus_tag="RP11-465B22.3-001"
/pseudo
CDS
join(2167..2290,2571..2857,3301..3492,3581..3680,4443..4566,5678..5758,6657..6737,6960..7094,7668..7815,8299..8446)
/locus_tag="RP11-465B22.3-001"
/notes="match: proteins: Q9QVA8 Sw:P29350 Sw:P35235 Sw:P41499 Sw:Q06124 Tr:AAP36054 Tr:BAC81774 Tr:BAC81775 Tr:P81718 Tr:Q64509 Tr:Q72M17 Tr:Q92124 Tr:Q96HD7 Tr:Q9CT18 Tr:Q9UK67"
/codon_start=1
gene
/product="protein tyrosine phosphatase family pseudogene"
join(2170..2292,4921..5547)
/locus_tag="RP11-465B22.1-001"
gene
join(AL645608.30:183584..183696,AL645608.30:184130..184243,2170..2292,4921..5547)
/locus_tag="RP11-465B22.1-005"
mRNA
join(2170..2292,4921..5547)
/locus_tag="RP11-465B22.1-001"
/product="novel transcript"
/notes="match: ESTs: Em:BF035622.1 Em:BI601978.1 match: cDNAs: Em:BC033949.2"
mRNA
join(AL645608.30:183584..183696,AL645608.30:184130..184243,2170..2292,4921..5547)
/locus_tag="RP11-465B22.1-001"
/product="novel transcript"
/notes="match: ESTs: Em:BI601978.1